DOCUMENT RESUME

ED 328 610 TM 016 141

AUTHOR Childs, Ruth Axman
TITLE Gender Bias and Fairness. ERIC Digest.
INSTITUTION American Institutes for Research, Washington, DC.; ERIC Clearinghouse on Tests, Measurement, and Evaluation, Washington, DC.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO EDO-TM-90-9
PUB DATE Dec 90
CONTRACT RI88062003
NOTE 4p.
PUB TYPE Information Analyses - ERIC Information Analysis Products (071)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Achievement Tests; *Aptitude Tests; College Entrance Examinations; Court Litigation; *Culture Fair Tests; Elementary Secondary Education; Legal Problems; *Sex Bias; *Test Bias; Test Interpretation; Test Results; Test Use
IDENTIFIERS ERIC Digests; New York; Scholastic Aptitude Test

ABSTRACT

A brief introduction to the topic of gender bias and fairness in testing is provided. A test is biased if men and women with the same ability levels tend to obtain different scores. The conditions under which a test is administered, the wording of individual test items, and a student's attitude toward the test can affect test results. While gender bias is a characteristic of the test itself, gender fairness refers to the ways test results are used. Detection of gender bias involves checking questions for: references that may be offensive to members of one gender; references to objects that are more familiar to men or to women; and unequal representation of men and women as actors or representation of members of each gender only in stereotyped roles. Gender bias assessments are particularly important for aptitude and achievement tests. Determining whether test results are being used fairly requires examination of the organizational policies that determine how test results are used. This latter issue is illustrated with a court case ("Sharif v. New York State Education Department") concerning the use of SAT scores as sole determinants of scholarship awards. A four-item list of additional reading materials is included. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED328610

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

GENDER BIAS AND FAIRNESS

by Ruth Axman Childs, ERIC/TM

TM016141

ERIC Clearinghouse on Tests, Measurement, and Evaluation December 1990 EDO-TM-90-9

Gender Bias and Fairness

by Ruth Axman Childs, ERIC/TM

Because test results are often the basis for decisions that affect students' educational futures, tests should provide equal opportunities for all students to demonstrate their abilities and knowledge. The issues of gender bias and fairness in testing are concerned with differences in opportunities for men and women.

This Digest provides a brief introduction to this complex topic. Commonly accepted definitions of gender bias and gender fairness are discussed. Approaches used to detect gender bias and fairness are introduced. Being aware of gender bias and fairness in testing will prepare you to intelligently question the uses of the test results which form the basis of decisions about individual test taker's futures.

What Is Gender Bias?

A test is biased if men and women with the same ability levels tend to obtain different scores. The conditions under which a test is administered, the wording of individual items, and even a student's attitude toward the test will affect test results. These factors may change with time as tests are administered differently, as items are revised, and as students feel more or less comfortable taking the test. The error caused by these factors will randomly affect both men and women.

Another type of error is caused by factors which do not change. Known as systematic error, it is the result of characteristics of the examinees that are stable (such as gender or race) and that are characteristics other than those the test is intended to measure. Gender bias in testing is often the result of such systematic error.

What Is Gender Fairness?

While bias is a characteristic of the test itself, fairness refers to the ways test results are used. Even an unbiased test may be used in ways that give an advantage to members of one gender. For example, a testing policy may treat test results differently for men and women.

Detecting Gender Bias in Testing

Test publishers go to great lengths to make sure that the questions contained in their tests are not biased and that the recommended uses of the tests are not likely to be unfair to members of one gender. However, no test is perfect. Careful examination of a test's questions may alert the educator to possible bias.

Test questions may be checked for:

- material or references that may be offensive to members of one gender,
- references to objects and ideas that are likely to be more familiar to men or to women,
- unequal representation of men and women as actors in test items or representation of members of each gender only in stereotyped roles.

If the questions involve objects and ideas that are more familiar or less offensive to members of one gender, then the test may be easier for individuals of that gender. Standards for achievement on such a test may be unfair to individuals of the gender that is less familiar with or more offended by the objects and ideas discussed, because it may be more difficult for such individuals to demonstrate their abilities or their knowledge of the material.

Although examination of the test items may reveal that the test contains questions which have the potential to yield biased results, such an examination may not be sufficient to determine bias. Statistical techniques are often used to examine for systematic gender differences.

Gender Bias in Aptitude Tests

Aptitude tests predict a future outcome. College entrance tests, for example, are usually administered to students during their senior year of high school and are intended to predict their grades during their first year of college. It would not be appropriate to check for bias in a college entrance test by just comparing average scores for men and women. You must also considering college grades.

American Institutes for Research ■ 3333 K Street NW. Suite 300 ■ Washington, DC 20007 ■ (202) 342-5060

TM016141

Determining whether such a test is biased involves using statistical techniques to calculate the predictive relationship separately for each gender. If the relationships are the same for men and women, we can say with confidence that the test predicts equally well for both genders.

These techniques have been used in recent studies concerning gender bias in college entrance tests. Several studies (such as those reported by Rosser in *The SAT Gender Gap*) have found that, while women tend to earn lower scores than men on some college entrance tests, they tend to have higher grade point averages during their first year of college. While inconclusive, these studies suggest that either 1) the predictive relationship between test scores and freshman GPAs are not be the same for both genders or 2) there is a systematic bias in the assignment of college grades.

Gender Bias in Achievement Tests

While aptitude tests predict future academic success, achievement tests measure an individual's knowledge. If it were possible to determine the exact amount of knowledge an individual possesses about a given topic, then establishing whether a test was biased would involve comparing the test scores with the actual knowledge levels and checking for systematic differences in these comparisons for men and women. Since it is not possible to obtain a perfectly accurate measure of an individual's level of knowledge, achievement test scores are usually compared with teacher-assigned grades and with scores from other tests. If one gender's scores on the test being considered are higher or lower than the scores of the other gender, but their grades in the classroom or their scores on similar tests are comparable for both genders, it may be that the test is biased.

Detecting Lack of Fairness in Testing

Determining whether test results are being used fairly requires the examination of the organizational policies that determine how the test results are used.

In a recent case, *Sharif* v. *New York State Education Department*, the plaintiffs charged that by using SAT scores as the sole basis for the award of state merit scholarships, the New York State Education Department was discriminating against girls who were competing for the awards. Although the girls tended to have higher high school grades than the boys competing for the scholarships, they also tended to have lower scores on the SAT, and so received fewer of the scholarships. The State Education Department argued that the SAT was the best objective measure available.

The SAT is intended to predict students' grades during the first year of college and does not claim to measure the achievement of students during high school. The stated intention of the New York scholarship program, however, was to base its awards on high school achievement. Since the program based its awards solely on the results of a test, the plaintiffs argued that the process denied girls a fair opportunity to demonstrate their eligibility for the awards. The court agreed and ruled that New York could no longer use SAT scores alone as a basis for these scholarship awards.

Not all educational policies have as easily measured and clearly different impacts on the genders as the policy described above. Considering whether such policies are supported by sound educational theory or research may be helpful in detecting possible difficulties.

Understanding Gender Bias and Lack of Fairness in Testing

The types of gender bias and lack of fairness described above may not affect all individuals equally. For example, although most girls have less experience than most boys with activities that help to develop mathematical skills, there are many girls who do not fit this generalization. Similarly, although most boys perform better than most girls on mathematics tests, some boys perform poorly on such tests.

Additional Reading

Campbell, P.B. (1989) *The Hidden Discriminator: Sex and Race Bias in Educational Research.* Groton, MA: Women's Educational Equity Act Program. ERIC Document Reproduction Service No. ED 322 174.

Klein, S.S. (Ed.) (1985) *Handbook for Achieving Sex Equity through Education.* Baltimore, MD: Johns Hopkins University Press. ERIC Document Reproduction Service No. ED 290 810.

Rosser, P. (1989) *The SAT Gender Gap: Identifying the Causes.* Washington, DC: Center for Women Policy Studies. ERIC Document Reproduction Service No. ED 311 087.

Tittle, C.K. (1979) *What to Do About Sex Bias in Testing.* Princeton, NJ: ERIC Clearinghouse on Tests, Measurement, and Evaluation. ERIC Document Reproduction Service No. ED 183 628.

Court Case

Sharif v. New York State Education Department, 709 F.Supp. 345 (S.D.N.Y. 1989).

This publication was prepared with funding from the Office of Educational Research and Improvement, U.S. Department of Education under contract number R-88-062003. The opinions expressed in this report do not necessarily reflect the position or policies of OERI or the Department of Education. ***Permission is granted to copy and distribute this ERIC/TM Digest.***